

BUNDESREPUBLIK DEUTSCHLAND

DE00/1944



REC'D	16 AUG 2000
WIPO	PCT

PRIORITY DOCUMENT
 SUBMITTED OR TRANSMITTED IN
 COMPLIANCE WITH
 RULE 17.1(a) OR (b)

**Prioritätsbescheinigung über die Einreichung
 einer Gebrauchsmusteranmeldung**

Aktenzeichen: 200 07 494.6

Anmeldetag: 26. April 2000

Anmelder/Inhaber: Thomas R o i t s c h, Regensburg/DE

Bezeichnung: Promotorsystem, dessen Herstellung und Verwendung

IPC: C 07 H, C 12 N

Die angehefteten Stücke sind eine richtige und genaue Wiedergabe der ursprünglichen Unterlagen dieser Gebrauchsmusteranmeldung.

München, den 18. Juli 2000
Deutsches Patent- und Markenamt
Der Präsident
 Im Auftrag

Faust



Promotorsystem, dessen Herstellung und Verwendung

Beschreibung

Die Erfindung bezieht sich auf die Genexpression und die Regulation von Genexpression in Pflanzen. Im Speziellen bezieht sich die Erfindung auf DNA-Promotor-Sequenzen, und auf Expressionskassetten, die in Pflanzen eingeführt werden können, um die Transkription einer benachbarten kodierenden Sequenz zeitlich und räumlich innerhalb der Pflanzen zu regulieren. Zusätzlich bezieht sich die Erfindung auf Expressionsvektoren, die solch eine Expressionskassette enthalten und die benutzt werden können, um Pflanzen zu transformieren.

Hintergrund zu der Erfindung

Ein Promotor ist eine DNA-Sequenz, die Expressionsort und -menge eines Genes beeinflußt oder bestimmt, und die Stellen für die Bindung der RNA-Polymerase zur Verfügung stellt. Die Position eines Promotors ist im Genom eines Organismus relativ zum Transkriptionsstartpunkt fixiert. RNA-Polymerase ist ein Enzym, das an den Promotor binden kann und die Transkription eines Genes vollzieht, das unter der Kontrolle dieses Promotors steht. Dabei entsteht die messenger-RNA, die wiederum zur Synthese des Proteins verwendet wird.

Promotoren sind in verschiedenen Organismen untersucht worden. Für bestimmte Spezies konnten konservierte DNA-Bereiche (sog. Consensus-Sequenzen) innerhalb von Promotoren gefunden werden, die mit verschiedenen Genen assoziiert sind. Von diesen Bereichen wird angenommen, daß sie in die Rolle, die der Promotor im Transkriptionsprozeß spielt, eingebunden sind.

Die Initiation des Transkriptionsprozesses in Pflanzen schließt eine Interaktion des Promotors mit der RNA-Polymerase II ein. Innerhalb von Pflanzenpromotoren wurden Consensus-Sequenzen oberhalb des 5'-Endes des Transkriptionsstartpunktes gefunden. Eine dieser Sequenzen ist etwa 7 Basenpaare lang und befindet sich etwa 20-30 Basenpaare oberhalb des Transkriptionsstartpunktes. Diese Sequenz ist als sog. TATA-box bekannt und es wird angenommen, daß sie eine Rolle bei der RNA-Polymerase Bindung spielt. Eine andere Sequenz mit einer Länge von ungefähr 9 Basenpaaren ist etwa 70-90 Basenpaare oberhalb des Transkriptionsstartpunktes zu finden. Diese Sequenz wird CAAT-box genannt und es wird angenommen, daß sie in der Regulation des Transkriptionslevels eine Rolle spielt. Es wurden

3

noch andere Regionen oberhalb des Transkriptionsstartpunktes identifiziert, die die Häufigkeit der Transkriptionsinitiation in Eukaryonten beeinflussen. Diese DNA-Bereiche, die Enhancer genannt werden, beeinflussen die Aktivität von Promotoren in ihrer Nachbarschaft. Diese Sequenzen sind jedoch der Definition nach keine Promotoren, da ihre Position nicht fixiert sein muß.

Um ein fremdes Gen in einem Organismus, z.B. einer Pflanze, exprimieren zu können, muß die kodierende Sequenz dieses Genes unter die Kontrolle eines Promotors gestellt werden und in die Pflanze eingebracht werden. Zur Insertion des zu exprimierenden Genes in das Pflanzengenom wird die fremde DNA meistens in das Ti-Plasmid von Agrobacterium tumefaciens gebracht, und dieses wird dann verwendet um die Pflanzen zu transformieren. Eine zweite häufig verwendete Methode ist die direkte Transformation von DNA z.B. mit Hilfe der sogenannten "particle gun". In den meisten Fällen werden hierfür bisher aus Bakterien isolierte Promotoren oder Promotoren von Pflanzenviren verwendet, die zur Expression des fremden Genes in den Pflanzen führen. Diese Promotoren haben den Nachteil, daß sie artfremd sind und daher den Kontrollmechanismen innerhalb der Pflanzen nicht unterliegen.

Bei Verwendung eines pflanzlichen Promotors ist die Expression eines fremden Genes möglich, die somit auch den pflanzlichen Kontrollmechanismen unterliegt. Durch Untersuchungen der Expression des Genes, vor dem der Promotor ursprünglich liegt, lassen sich genaue Kenntnisse über die Expressionsstärke, die Zeit, zu der das Gen exprimiert wird, und den Expressionsort sammeln, die auf die Expression eines fremden Genes, das unter die Kontrolle dieses Promotors gestellt wird, weitgehend übertragbar sind. Ein weiterer Vorteil ist, daß bei Verwendung eines genau charakterisierten, pflanzlichen Promotors gezielte Eingriffe und Untersuchungen in die Entwicklung bestimmter Pflanzenteile möglich ist.

Im vorliegenden Fall wurde der Promotor einer Invertase aus Tabak kloniert, die in sehr großen Mengen, aber nur in einem bestimmten Entwicklungsstadium und hochspezifisch, nur in den Antheren exprimiert wird. Zur Klonierung des Promotors wurde eine genomische Bank aus Tabak mit einer Sonde aus dem kodierenden Bereich der untersuchten Invertase durchsucht. Die erhaltenen genomicischen DNA-Sequenzen wurden mit molekularbiologischen Methoden weiter charakterisiert und schließlich ein Klon isoliert, der die Promotorsequenz enthielt, die dann sequenziert wurde. Es wurden Expressionskassetten mit verschiedenen Genen hergestellt, die in Agrobakterium tumefaciens eingebracht wurden, um Tabakpflanzen damit zu transformieren. Dieses neue Promotorsystem kann nun dazu verwendet werden, fremde Gene in Pflanzen in

4

einem bestimmten räumlichen und zeitlichen Rahmen zu exprimieren. Das Wort fremd meint dabei Gene, die nicht natürlich in Verbindung mit diesem Promotor vorkommen. Außerdem kann mit Hilfe dieses Promotorsystems die Expression von fremden und eigenen Genen in der Pflanze moduliert werden, d.h. Gene können überexprimiert oder reprimiert werden.

5

Ausführungsbeispiele:

1. Die Promotor-DNA-Sequenz, bestehend aus den 4300bp des 5' liegenden DNA-Bereiches, gezählt oberhalb vom Translationsstartpunkt der antherenspezifisch exprimierten Invertase aus Tabak, oder Teile davon, die dadurch gekennzeichnet sind, antherenspezifische Expression für dahinter gelegene Gene zu vermitteln, kann mit Hilfe molekularbiologischer Methoden so verändert werden, daß Restriktionsschnittstellen am 5'- oder 3'-Ende eingefügt werden.
2. Die wie in 1. veränderte Promotor-DNA-Sequenz oder Teile der Promotor-DNA-Sequenz können zur Herstellung einer Expressionskassette zur Expression fremder Gene in Pflanzen verwendet werden. So eine Expressionskassette ist dadurch gekennzeichnet, daß sie a) die Promotor DNA oder Teile davon aus 1. enthält; b) eine Verbindungs-DNA ohne spezielle Funktion bzw. fremde Gene, verbunden mit der ersten Schnittstelle beinhaltet; und c) eine 3'-Region, bestehend aus der 3'-Region eines eukaryontischen Genes enthält, wobei diese 3'-Region eine zweite Restriktionsschnittstelle an ihrem 5'-Ende besitzt und diese 3'-Region über diese zweite Restriktionsschnittstelle mit der Verbindungs-DNA bzw. den fremden Genen aus b) verbunden ist.
3. Eine Expressionskassette wie in 2. beschrieben, kann in einen Expressionsvektor kloniert werden. Dieser Expressionsvektor kann dazu verwendet werden, mit Hilfe verschiedener gängiger Methoden (z.B. Agrobakterien vermittelte Transformation; direkte Transformation) Pflanzen zu transformieren. Transgene Pflanzen können dann unter Bedingungen angezogen werden, unter denen das fremde Gen unter der transkriptionellen Kontrolle der beschriebenen Promotor-DNA-Sequenz exprimiert wird.
4. Die Promotor-DNA oder Teile davon können dazu verwendet werden, die Translation eines pflanzen-eigenen Genes zu modulieren. So kann die Expression eines Genes durch Einbringen weiterer Kopien unter der Kontrolle dieses Promotors gesteigert werden oder die Expression kann mit Hilfe der Antisense-Technik unterdrückt werden. Dazu muß die DNA-Sequenz des zu unterdrückenden Genes in "verkehrter Richtung" in eine Expressionskassette wie unter 2. beschrieben kloniert werden und wie unter 3. beschrieben in Pflanzen transformiert werden.
5. Die spezifischen Eigenschaften der Promotor-DNA-Sequenz bzw. von Teilen davon, ermöglichen in transgenen Pflanzen, die wie unter 3. beschrieben hergestellt wurden, eine

zeitlich (nur während der Pollenbildung) und räumlich (nur in Antheren) definierte Expression von fremden Genen in Pflanzen.

6. Aufgrund des starken Expressionslevels der antherenspezifischen Invertase in Tabak, lassen sich mit Hilfe dieser Promotor-DNA-Sequenz bzw. mit Teilen davon, in transgenen Pflanzen große Mengen eines bestimmten Proteins zu einem bestimmten Zeitpunkt und in einem bestimmten Ort der Pflanze (siehe 5.) herstellen. Dieses Protein kann dann durch Ernten der Antheren, Aufschluß und für das hergestellte Protein spezifische Reinigungsverfahren in großen Mengen gewonnen werden.

7. Die Promotor-DNA-Sequenz, oder Teile davon, können dazu verwendet werden, in die Entwicklung der Antheren in Pflanzen einzutragen. So können transgene Pflanzen hergestellt werden, bei denen beispielsweise durch Antisense-Expression von Invertasesequenzen die Proteinmengen für die extrazelluläre Invertase verringert werden. Dies führt zu männlich sterilen Pflanzen, die in der Landwirtschaft, bei der Herstellung von Hybridsaatgut von großer Bedeutung sind.

8. Die Verwendung des Promotorsystems zur Herstellung männlich steriler Pflanzen, wie z.B. unter 7. beschrieben kann als Sicherheitssystem bei der Herstellung anderer, kommerziell oder wissenschaftlich nutzbarer transgener Pflanzen verwendet werden. So besteht bei Verwendung männlich steriler Pflanzen nicht die Gefahr des Auskreuzens der genetischen Veränderungen auf Pflanzen die auf benachbarten Feldern oder wild wachsen. Die begrenzte Natur des Eingriffes der zur männlichen Sterilität bei Verwendung dieses Promotorsystems führt stellt einen besonderen Vorteil dar, da nicht in das vegetative Wachstum der Pflanze eingegriffen wird und keine Wechselwirkungen mit den zusätzlich eingebrachten genetischen Veränderungen zu erwarten sind.

9. Die Promotor-DNA-Sequenz, oder Teile davon, können dazu verwendet werden, transgene Pflanzen herzustellen, die pflanzeneigene Stoffe in großer Menge herstellen, die positiv auf ihre Entwicklung, insbesonders betreffend den Ertrag von fruchttragenden Pflanzen, wirken können. Beispiele für solche pflanzeneigene Stoffe wären Wachstumshormone oder Proteine die zur Energieversorgung der wachsenden Gewebe (z.B. Invertasen, Zuckertransporter) notwendig sind.

10. Mit Hilfe des Promotorsystems können transgene Pflanzen hergestellt werden, in denen die Menge der produzierten pflanzeneigenen Stoffe (z.B. Wachstumshormone) reduziert werden kann. Diese Reduktion kann durch Einbringen von abbauenden Enzymen, von Inhibitoren oder durch sog. "single-chain" Antikörper erreicht werden.

11. Um Erträge von männlich sterilen, fruchttragenden Pflanzen zu erhalten und zur Vermehrung männlich steriler Pflanzen, die unter Verwendung des Promotorsystems wie z.B. unter 7. beschrieben hergestellt wurden können Restorerstämme hergestellt werden, die nach einer Kreuzung mit den männlich sterilen Stämmen zu fertilen Pflanzen in der F1-Generation führen. Restorerstämme für die unter 7. beschriebenen männlich sterilen Pflanzen könnten Proteine enthalten, die die Kohlenhydratversorgung der Antheren wieder herstellen können, wie z.B. artfremde Invertasen (z.B. aus *Saccharomyces cerevisiae* oder Bakterien), oder Saccharosetransporter in Verbindung mit intrazellulären saccharosespaltenden Enzymen (z.B. Saccharose Synthase, neutrale oder vakuolare Invertasen).

12. Alternativ zur Herstellung von Restorerstämmen wie unter 11. könnten Pollen von männlich sterilen Pflanzen in einer *in vitro* Kultivierung zu fertilen Pollen entwickelt werden, mit denen dann eine Befruchtung der transgenen Pflanzen stattfinden kann.

13. Pollen können durch eine *in vitro* Embryogenese zu haploiden Pflanzen herangezogen werden. Diese haploiden Pflanzen können dann zu homozygoten diploiden Pflanzen gezüchtet werden. Zur Induktion dieser *in vitro* Embryogenese ist ein Hunger- und Stress-Schritt nötig. Da transgene Pflanzen, die unter Verwendung des Promotorsystems wie z.B. unter 7. beschrieben hergestellt wurden, in ihrer Zuckerversorgung gestört sind, sind die Pollen dieser Pflanzen bereits in Richtung Embryogenese determiniert und es bedarf keiner zusätzlichen Hunger- oder Stress Behandlung mehr. Die Embryogenese läuft in diesen Pflanzen daher schneller und effizienter ab.

Erreichte Vorteile:

Mit Verwendung dieses Promotorsystems steht ein Werkzeug zur Verfügung, mit dem man zeitlich und örtlich gezielt fremde Proteine in Pflanzen exprimieren kann und pflanzeneigene Gene in ihrem Expressionslevel modulieren kann.

Ansprüche

1. Promotorsystem, gekennzeichnet durch einen Klon aus genomicischen DNA-Sequenzen.
2. Promotorsystem, gekennzeichnet durch Expressionskassetten.
3. Verwendung eines Promotorsystems nach Anspruch 1 oder 2 zum Beeinflussen eines Genes.
4. Verfahren zur Herstellung eines Promotorsystems nach Anspruch 1 oder 2.
5. Verfahren zum Experimentieren mit einem Promotorsystem nach Anspruch 1 oder 2.

Zeichnung 1: Promotor-DNA-Sequenz der extrazellulären Invertase aus Tabak

1 TCTAGAACATGA CGCCACCGGCC CAGGACGGGG AGTATGATT CCCCGAATGT
 51 TCGTTCAACT GCATTGTTAA AACCTGTTAG CGTGATGCAG CCCGGTACTA
 101 TCTTATCCTC GAGTTCATT TGTGCAAGTA CTCGAGGATG GACAATTAC
 151 GGGCCACTCC CATCGTCCAC CATAATGCGT CTTACATCTG TATCTAATAT
 201 TCGTAAAGTG ATAACGAGGG CATCATAGTG AGGGAAAACC AAACCGTGGT
 251 TATCTGACTT ATCGAAGATG ATACTTTCTT TAAGTTTCTC GTACCGTTCA
 301 TGAGTGATTA ACTGTTGAG CTTGTTGGTT GTGGCGAACT TTACGTTGTT
 351 GATCGAAACG TCGTCTCCGC CCCCAGTGT AATGTGAATG GTGCCAGTCG
 401 GTAAGGGTGG TTTCGCGGT CCCTGGTGGT GTTACAGTCC TCGAGAAAAG
 451 TTGGTCCTTC CTGGTCACA CAACAAATATT TTGAGGTGTC CTTGATGAAG
 501 CATGTCCATG ACCCTCTTGTGTC TTAGGGCGAT ACAATCCTCA GTTTGTCGAC
 551 CTCGCTCTTG GTGGAACTCG CAGAGGGCAT CTGATTTCT AGTGCTTGG
 601 TCTGACCTCA TCTTTTGTGG CCACTTTACT TTTGGTCCGA GCTTCTTCAA
 651 TGCATAGACT ATTTCTGAGG GTGACACACA AAATTGTGA GCGGATAGTA
 701 AAGAGGGCAT ACCTCTCTCG TTCCGGTGAG TCCCTGTCCT TGGCCTAGAT
 751 GGGCCCTCTT CGTAGCGGGA GAGGGGCATG ATGGCACTTT TGACATATGG
 801 TTGATCCATT TCTCGGTTAG ATCATGGAGC TGCAAGATCT CTCTTGGCAT
 851 CATTGGACG ATCCCTCCTG GTTTGGCTT GTACCGAGGT CAATCGATGA
 901 GTTGGCCAT TCAGGTGTC TCAGTCCGGCA CGGGCCTCAG CACAGTAGGC
 951 GTTGTGTATT TCATCCAAG TGGTGGACG ATATTCATA AGTTGGTTA
 1001 ACAGTTTCT GGTCGCCCTC GAGCCATTCA TGTTCAGCCC ATTCTGGAA
 1051 GTTGCTACAA CCATTCCTTC TGATACATTC GGTAAGGTCA TCCTTACTCT
 1101 GTTGAATCGA GCGAGGAAGT CCCTCAATCC CTCTCCGAGT GATTGTTGA
 1151 TGGCAAATAT ATCGTCACG CTTGCTCCG CGTTTTAGC CCCAACATGG
 1201 GCCATTATGA ACTTGTGCGC CATCTCTTCG AATATTCAA TGGAGGCCGC
 1251 GGGCAGCTGT GAATACCAAG TCAATGCTCC TCCGGTAAGG GTCTCGCCGA
 1301 ACATTTCAA CAAGATGGAG GAGACTTGTG CTTTGGAGAG ATCATTGCC
 1351 TTGACATAATG ATTACATGAT CTTGGGGTC GGTCGTACCA
 1401 TCATAAATTG TCAGATAAGG TGGCATCTG AACGTCTTGG GTATGGCATA
 1451 TGGGGCGGCT TCATCACTGT AGGGTTGCTC GACTAACCGA CCAGCGTCTC
 1501 TTTTGGAAA TATTTTGGG GCAACCGGTA TTTTATCGAC TCTTTCTTGG
 1551 TGTTCTCTCA TTGATCCCCG AAGCATTITA TTTTCGTTT CCATTTCTTC
 1601 CATTTCCTTC AGAATGGCCG TGAGGGTGTG ATTACCTGCA TTATTAATAT
 1651 TGTGAGTGT ACCTGTTACT GAAGGGGGAG GTGCGTGTG TTTGGTCATT
 1701 GCTGGTCAA TGCAAGTCCT TGCATTTCT CTAAATACCT CCTGAGTGGG
 1751 TTTGTTGAGG ATGCCGGTCA GCAATTTGT CAGCCAAGCT TCGAGTAGCT
 1801 TCTTCACCGC TGGTGGCGCC TCTTCCGTTG TGGCACGTGGA AGCTCCTTAA
 1851 CCGCGGGATG TTGCGATACT GCTGTGAGGG AGGGGTGATC CACTTCGTGG
 1901 GGGAGAGGTG TTAGCCGTTA TGCCTTCCTT TTCTATTCG GAGACCTCAT
 1951 TGATGGTGT TAAGAGGTG GTAGTGAGAT TGGCCACTGC CTTCATCCTT
 2001 TCTTCTCCCT TACCTGCCAT GTCAGATCTG GGTGTACAAG GAAGTAGGAG
 2051 CTTCTCTTCG TCTTTTTGT GAAATGTGCC AGTTATAGAT CTAAAAGAAA
 2101 CTAAAGTTT AACCTGACTA TCCCTCACAGA CGGCGCCAAA TTGTTGACC
 2151 AAAAAATATA GACTTTGTG TAAATTAAATT AATATTGTAT GACAAAGGAT
 2201 TAAACCTAGT TAATGATAAT AACTTCAGAT CTATAATCAA TTAACAGCAA
 2251 TCACGGTCAT AGCAGCGTTG AGAGAAGATT AAATGTGATG TYCATTCAAT
 2301 ATTTCAAGAT CATTAAATGAT AGGGGAATAT CAAGCAATAA ATAACGATAA
 2351 ATGGCATTAAG AGTAAATAAG GAGAATGATT CACCAATAT TGAATGAGGT
 2401 GGATGATTCT TCTTTTGAC AATGATGAAT GATGGGCAA TACTAGAATG
 2451 TTGGGACCCCT TCTCGGATCT AATGAAAAAA GTATGGAATA GTAGATAATC
 2501 GAAATCTCTT AGAAAGGTAG TGATTGTCTT TTATCTAGAG AGAAAGTCTG

10

2551 CTTTTCAGAAG AATATTTTA TCAAGAGAATA TTACATCCCC CTCTCTCCCT
2601 ATCTCTTTTTCTATTTATAT GGGACATTCG TGAATCAATC CTAAAAGTAC
2651 ATACACCAAG AATATTCAT AAAATAITTT TTGAAATATT CTATTTATAAA
2701 AACTAGCIGT TAGCACTCGA CCTCGGTGCGY TATTGACTAC TCGGTTACGA
2751 GCCCTGTCAT TTACTAATCG ACCTCGATTA CATCACTTTC TACGATACTG
2801 CTTCATGTCA AATCTTAATG AAAGCAGATT TTGACCCATA CAATAATATG
2851 ACAAAATTGC TTCCAAAGAA AACATGGCTC TTATAGTGAA ATATCGTTAG
2901 ACTGTTATAG AAAGATCTGA ATTATTTAT AAGAATAGTG TTTTTTCTT
2951 TTCTTTCAT ATCTAAGGAG TAAAGCAACC ATGAATAGAA AAGGCTTAGT
3001 AACTATATAT CAAAGGAATG GTGTTTTTC TTTAAATATG GATAAAAATT
3051 TGTGAATATA CAAGATTAGA TCAATTAACA AAGGTTATGG TGGAGTGGTA
3101 AGCAGAGGCCG GACCTATGTG TTATAGTAAG GGGTCACCCA CTACTAGAAA
3151 TCCGGTAAAG ATCGATCAA AAACCGAACCA ACATTGGTCG GTAATGGCCA
3201 AAAACTGACC AAAACCGCAT CATTACGTG TGAACGGTAT TTTATGGTC
3251 GGAAAGGAAT ACCGACCAA CTGGTGGGA AATTACCGAC CAACTTTGGT
3301 CGGTCAATTA AATTCAAAAA AAATATTGTA AAAAAAAACC GACCAAAGTT
3351 GATCGGTATT TTAAATTATGT AATLAAAAGA TTCACTATCT GGGAAATCGAA
3401 CCGGGGTCTG TACTATGGCA AGATACTATT CTACCACTAG ACCATTGGTT
3451 CATTGGTGTG TAAGACTGTC TTTTATTGTA TTTATACCTCT TTAATTATAT
3501 TTTTGACCGA AAATAACCGA CCAAAGTTGG TCAGTTTAT TAAAAGTAA
3551 AATTACTTAC CAAAGTTGGT CGATTTTTT AAATGATCCG CCGAATTAAAC
3601 CGACCAATT TGTTAGGTT TTTAAATATT AATTTTTATT TATTTAATT
3651 GAAAAAACTAA CCAAAGTTAG TCGGTTCTT GAAACATAAA TTTCGCGGGA
3701 CTCAAAAATA GTTCCCGCA TTTTGCGCC AAAGAAAACC GACCAAAGTT
3751 GGTGGTTTC GTAAAAAAA AAAAAAATTAA AAAAAATATAT TTTAAAANAC
3801 CGACCAACTT TAGTCGGTTT TTGGTCGAT TTTTGACCG ACCAAAGTTG
3851 GTCGGTGCAC CTTGGTCGGT TTTGCGGAA TTTCTAGTAG TGACCGAAC
3901 CTGTAAGCTT CGGGAGAAAT TTGTTATATG TATATGTGTA TATCTTAAA
3951 ATGATTAATT TAAAGAACGT GGCACCCCTGA ATACTAGAAG CCTTTAGGGG
4001 CACTAGATGA GCAGAATAAC GTGTTCTCGT CGCGTAAAAA TACTTGGATC
4051 CGCCTATGAT GGTAAGTACT TCTTCGTCT TAATCAGAGG TTTCGACTTC
4101 GAGCTCCAGA TATAAACTAT AGACTCGTCT TTATAGCACC TTTTAATAAG
4151 ACTATGACTT CATCTGATTT CTCTATAAAT ACTCCTCAAG CTTTCGGTTC
4201 TTCTCCATTC TTCAGTTCT TTCTCCACAT CACAGAAGTG AAAACAAAAC
4251 AAGAAGAAGA AGAAGAAGAA AAATAAAGAG TTCTGTCAA ATTAAGTCCA
4301 ATAGGGAAAA TG